nature ecology & evolution

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41559-019-1039-3

In the format provided by the authors and unedited.

The social coevolution hypothesis for the origin of enzymatic cooperation

Samuel R. Levin[1]*, Sylvain Gandon[2] and Stuart A. West[1]

[1]Department of Zoology, University of Oxford, Oxford, UK. [2]CEFE UMR 5175, CNRS, Université de Montpellier, Montpellier, France.
*e-mail: samuel.r.levin@gmail.com

Supplementary Material for "The social coevolution hypothesis for the origin of enzymatic cooperation"

Samuel R. Levin[1], Sylvain Gandon[2], and Stuart A. West[1]

[1]Department of Zoology, University of Oxford, UK
[2]Centre d'Ecologie Fonctionnelle et Evolutive (CEFE), France

S1 The Model

We need a total of seven equations to capture the ecological dynamics of both the residents and the mutants. From the life cycle in figure one and the description in the main text, this allows us to write (for the residents):

$$\begin{aligned}
\frac{d[X]}{dt} =& \left(\left(\left(r_X\left(1-d\right)\eta\right)-\mu_X\right)[X]\right) \\
&-\left(\beta\left(1+\zeta c\right)[X][Y]\right) \\
&\quad+\left(\left(\left(1+\kappa\right)\left(1-d\right)r_X\eta\right)+\mu_Y+\left(1-\xi c\right)\delta\right)[XY] \\
\frac{d[Y]}{dt} =& \left(\left(\left(r_Y\left(1-c\right)\eta\right)-\mu_Y\right)[Y]\right) \\
&-\left(\beta\left(1+\zeta c\right)[Y][X]\right) \\
&\quad+\left(\left(\left(1+\kappa\right)r_Y\left(1-c\right)\left(1+\omega d\right)\eta\right)+\mu_X+\left(1-\xi c\right)\delta\right)[XY] \\
\frac{d[XY]}{dt} =& \beta\left(1+\zeta c\right)[Y][X] \\
&-\left(\mu_Y+\mu_X+\left(1-\xi c\right)\delta-\left(\left(1+\lambda d\right)\left(1+\alpha c\right)\right)r_{XY}\eta\right)[XY]
\end{aligned} \tag{S.1}$$

Where $\eta = 1 - k([T] = [X] + [XY] + [Y])$ is density dependent regulation, and k controls its extent.

Following the standard adaptive dynamics approach, we assume that invading mutants are rare enough so as not to affect the dynamics of the resident population (Metz et al., 1992; Rand et al., 1994; Dieckmann and Law, 1996). Accordingly, we only need four additional equations to capture the dynamics of mutants in each gene (X' and Y'), which are expressions for $\frac{d[X']}{dt}$, $\frac{d[Y']}{dt}$, $\frac{d[X'Y]}{dt}$,

and $\frac{d[XY']}{dt}$. These differ from system S.1 only in their values for c and d, the mutant trait values, which we denote with primes as c' and d'.

The equations for the mutant can be written in the form (Van Baalen and Jansen, 2001):

$$\frac{d[i']}{dt} = F_i'[i'] - \beta'[i']\overline{[j]} + P_i'[i'j] \tag{S.2}$$

$$\frac{d[i'j]}{dt} = \beta'[i']\overline{[j]} - M_{ij}'[i'j] \tag{S.3}$$

Where $F_i = (\rho_i - \mu_i)$ is the growth of i alone, $P_i = (\mu_j + \theta_i + \delta)$ is the production of i's from complexes, and $M_{ij} = (\mu_i + \mu_j + \delta - \rho_{ij})$ is the loss of complexes. The primes indicate mutant values in gene i, and $\overline{[j]}$ is the equilibrium frequency of copies produced in the absence of the mutant. This form for an invasion condition was first identified by Van Baalen and Jansen (2001). For illustration, we reproduce each term for replicator Y, but equivalent equations can be extracted for replicator X.

$$\begin{aligned}
F_Y' &= (r_Y (1 - c') ((1 - k([T])))) - \mu_Y \\
P_Y' &= ((1 + \kappa) r_Y (1 - c') (1 + \omega d) ((1 - k([T]))) + \mu_X + (1 - \xi c') \delta) \\
M_{XY}' &= (\mu_Y + \mu_X + (1 - \xi c') \delta - ((1 + \lambda d) (1 + \alpha c')) r_{XY} ((1 - k([T])))) \\
\beta_Y &= \beta (1 + \zeta c')
\end{aligned} \tag{S.4}$$

We can rewrite system S.4 in matrix form as:

$$\frac{d}{dt}\begin{bmatrix}[i'] \\ [i'j]\end{bmatrix} = \begin{bmatrix} F_i' - \beta'\overline{[j]} & P_i' \\ \beta'\overline{[j]} & -M_{ij}' \end{bmatrix}\begin{bmatrix}[i'] \\ [i'j]\end{bmatrix} \tag{S.5}$$

The first matrix on the right hand side of equation S.5 contains all the information we need to determine the spread of a rare mutant (Van Baalen and Jansen, 2001; Hurford et al., 2009). A useful decomposition of this matrix is the form $\mathbf{F_i}$ - $\mathbf{V_i}$, where,

$$\mathbf{F_i} = \begin{bmatrix} F_i' & P_i' \\ 0 & 0 \end{bmatrix}, \mathbf{V_i} = \begin{bmatrix} \beta_i'\overline{[j]} & 0 \\ -\beta_i'\overline{[j]} & M_{ij}' \end{bmatrix} \tag{S.6}$$

According to the next generation theorem (Hurford et al., 2009), given that $\mathbf{F} > 0$, $\mathbf{V^{-1}}$, and the spectral bound of $-\mathbf{V}$ is negative, the condition for a mutant to invade is that the spectral radius of $\mathbf{F}\mathbf{V}^{-1} > 1$. This condition can be written as:

$$\frac{P_i'}{M_{ij}'} + \frac{F_i'}{\beta'\overline{[j]}} > 1 \tag{S.7}$$

This is equation 2 in the main text.

Taking the derivatives of equation S.7 with respect to small changes in mutant trait values gives the direction of selection in each trait, which we use to produce Figure 2c.

Coevolution is driven by the interaction between the two mutant trait values, c' and d'. These interaction terms are contained entirely in the derivative of the first term of Equation (S.7, and it can easily be seen that $\frac{\partial}{\partial c'}\left(\frac{P_i'}{M_{ij}'}\right)$ and $\frac{\partial}{\partial d'}\left(\frac{P_i'}{M_{ij}'}\right)$ are positive functions of c' and d', respectively.

This is not a quantitative model, and therefore we do not extensively discuss the specific values the parameters can take. Those with a specific chemical system in mind should use parameter values relevant to the biochemistry of the molecules at hand. Here, we simply point out that what matters for the evolution of cooperation is the ratio of different key parameters, found in equation (S.7). For example, if the baseline replication rate, r_i, is of several orders smaller than the baseline association rate, β, *and* the benefit to cooperation, ω, is small relative the other parameters, it will be difficult for cooperation to spread. This is because in the absence of an association mutation, individuals often already find themselves in pairs, and the additional benefit of cooperation is insignificant. All the important relationships can be read directly from equation (S.7) (equation 2 in the main text).

S2 Tracking same-type replicator pairs

Above we did not track XX or YY pairs. This means that the above model holds in systems where the replicators do not form self-self complexes. We also conjectured that the results would approximately hold even if they do form such complexes, because individuals in XX and YY pairs do not gain byproduct benefits, and therefore have a lower replication rate than when in XY complexes. We checked this by developing a model that explicitly tracks such pairings. This requires two additional equations for the density of XX and YY complexes, for a total of five equations:

$$
\begin{aligned}
\frac{d[X]}{dt} =& \left(\left(r_X\left(1-d\right)\eta\right)-\mu_X\right)[X]\right) \\
&-\left(\beta\left(1+\zeta c\right)[X][Y]\right) \\
&\quad+\left(\left(\left(1+\kappa\right)\left(1-d\right)r_X\eta\right)+\mu_Y+\left(1-\xi c\right)\delta\right)[XY] \\
&\quad-\beta[X][X] \\
&\quad+\left(\mu_X+\delta+r_X\left(1-d\right)\eta\right)[XX] \qquad \text{(S.8)} \\
\frac{d[Y]}{dt} =& \left(\left(r_Y\left(1-c\right)\eta\right)-\mu_Y\right)[Y]\right) \\
&-\left(\beta\left(1+\zeta c\right)[Y][X]\right) \\
&\quad+\left(\left(\left(1+\kappa\right)r_Y\left(1-c\right)\left(1+\omega d\right)\eta\right)+\mu_X+\left(1-\xi c\right)\delta\right)[XY] \\
&\quad-\beta[Y][Y] \\
&\quad+\left(\mu_Y+\delta+r_Y\left(1-d\right)\eta\right)[YY] \qquad \text{(S.9)} \\
\frac{d[XY]}{dt} =& \beta\left(1+\zeta c\right)[Y][X] \\
&\quad-\left(\mu_Y+\mu_X+\left(1-\xi c\right)\delta-\left(\left(1+\lambda d\right)\left(1+\alpha c\right)\right)r_{XY}\left(1-k\left([T]\right)\right)\right)[XY] \\
\frac{d[XX]}{dt} =& \beta[X][X]-\left(\mu_X+\mu_Y+\delta-r_{XX}\eta\right)[XX] \\
\frac{d[YY]}{dt} =& \beta[Y][Y]-\left(\mu_Y+\mu_X+\delta-r_{YY}\eta\right)[YY]
\end{aligned}
$$

Where $\eta = 1-k([X]+[XY]+[Y]+[XX]+[YY])$.

Following the same approach as above, we derive the condition for a mutant in replicator type i to spread as:

$$
\frac{\overline{[i]}\beta}{\overline{[i]}\beta+\overline{[j]}\beta'}\left(\frac{P'_{ii}}{M'_{ii}}\right)+\frac{\overline{[j]}\beta'}{\overline{[i]}\beta+\overline{[j]}\beta'}\left(\frac{P'_{i}}{M'_{ij}}+\frac{F'_{i}}{\overline{[j]}\beta'}\right)>1 \qquad \text{(S.10)}
$$

The new terms, P'_{ii} and M'_{ii}, capture the production and loss of same type pairs, respectively. This inequality is of a similar form to Equation (S.7). The original expression for fitness is now weighted by the relative rate of pairing with the other type. The new component of fitness (the first term) the ratio of production of same type pairs to loss of same type pairs, and is weighted by the relative rate of pairing with the same type.

Numerically solving across parameter state space shows that the same results hold as above, with cooperative enzymatic activity failing to spread on its own, association evolving in the absence of such activity, and the two traits co-evolving to higher values than when on their own (Figure S1).

S3 An explicit model of pairing

The above model left unspecified how cooperation and association increase pairing of replicator copies, capturing the effect in the term ρ_{ij}. We now adapt the

model to a specific population structure, in order to make this effect explicit. Doing so necessarily requires sacrificing some of the generality of the first model, but what it loses in generality it gains in precision.

We now need to track two additional populations: X's and Y's that have been produced from pairs. This is because in order to explicitly model the effects of cooperation and association on pairing, we need to track the densities of copies produced from pairs before they become randomly mixed in the population.

The assumptions are the same as above, except now we allow for some baseline rate, χ, at which copies produced from pairs immediately pair again. Otherwise they return to the independent populations of X and Y. We assume that the rate of pairing is increased by both cooperation and physical association, by a factor $(1+\lambda d')(1+\alpha c')$, and is a function of the densities of copies produced, denoted $[X_o]$ and $[Y_o]$. The new system of equations describing the population dynamics is now:

$$\begin{aligned}
\frac{d[X]}{dt} =& (((r_X(1-d)\eta) - \mu_X)[X]) \\
& - (\beta(1+\zeta c)[X][Y]) + \psi[X_o] \\
\frac{d[Y]}{dt} =& (((r_Y(1-c)\eta) - \mu_Y)[Y]) \\
& - (\beta(1+\zeta c)[Y][X]) + \psi[Y_o] \\
\frac{d[XY]}{dt} =& \beta(1+\zeta c)[Y][X] \\
& - (\mu_Y + \mu_b + (1-\xi c)\delta)[XY] + (1+\lambda d)(1+\alpha c')[X_o][Y_o] \\
\frac{d[X_o]}{dt} =& ((1+\kappa)(1-d)r_X\eta)[XY] - (1+\lambda d)(1+\alpha c')[X_o][Y_o] - \psi[X_o] \\
\frac{d[Y_o]}{dt} =& ((1+\kappa)r_Y(1-c)(1+\omega d)\eta)[XY] \\
& - (1+\lambda d)(1+\alpha c')[X_o][Y_o] - \psi[Y_o],
\end{aligned} \tag{S.11}$$

where $\eta = 1 - k([X] + [XY] + [Y] + [X_o] + [Y_o])$. The parameter ψ controls the relative rate at which copies produced from pairs return to the population of free X and Y.

Following the same approach as before, we can write system S.11 in matrix form as:

$$\frac{d}{dt}\begin{bmatrix}[i'] \\ [i'j] \\ [i'_o]\end{bmatrix} = \begin{bmatrix}F'_i - \beta'\overline{[j]} & P'_i & \psi \\ \beta'\overline{[j]} & -M'_{ij} & A'_i\overline{[j_o]} \\ 0 & PR'_i & -A'_i\overline{[j_o]} - \psi\end{bmatrix}\begin{bmatrix}[i'] \\ [i'j] \\ [i'_o]\end{bmatrix} \tag{S.12}$$

$P'i$ now measures only replicators returned to the independent population from complexes as a result of dissociation and destruction, because copies produced from complexes are captured in the term PR'_i. A'_i measures the association of copies produced from complexes, and $\overline{[j_o]}$ is the equilibrium frequency of copies produced from complexes.

Using the next generation theorem (Hurford et al., 2009), we find the condition for a mutant to invade a resident population to be:

$$\frac{\psi P_i' + \overline{[j_o]} P_i' A_i' + \psi P R_i'}{\psi M_{ij}' + \overline{[j_o]} M_{ij}' A_i' - \overline{[j_o]} P R_i' A_i'} + \frac{F_i'}{\beta' \overline{[j]}} > 1 \tag{S.13}$$

This equation and its derivatives with respect to changes in c' and d' allow us to analyse evolution of cooperation or physical association on their own, or their co-evolution. We recover the result that co-evolution can favour the evolution of cooperation in conditions under which it would not have evolved on its own (Figure S2). The result depends crucially on the relative rate at which copies produced from pairs return to the independent populations (ψ), with $\psi \approx 1$ recovering the main results.

S4 Relation to previous mathematical models

Adaptive dynamics was developed through a serious of papers in the 1990s (Metz et al., 1992; Rand et al., 1994; Geritz et al., 1997; Dieckmann and Law, 1996). The key assumption is that the mutant is initially rare enough that you can assume it does not impact the ecological equilibria. Accordingly you can assume that the resident populations have reached equilibrium when the mutant is introduced, and study it's growth rate in that setting.

Law and Dieckmann (1998) developed a model to study the coevolution of two species in the context of an exploiter-victim relationship evolving into a vertically transmitted symbiosis. They used the same approach of tracking both independent populations as well as complexes (in their case 'holobiont'). Their main result was that, when costs of being free-living are high enough, even strictly exploitative relationships can evolve into symbioses in the presence of vertical transmission. This is analagous to our result that, when benefits of being in a complex are high enough, stickiness can favour the evolution of cooperation.

Van Baalen and Jansen (2001) extended the work of Law and Dieckmann (1998) in developing a model to study two-species systems, in which they tracked the two independent populations as well as complexes. Their goal was to develop a general methodology for studying two interacting populations, and they specifically discussed two prey populations which shared defence of a predator and a host parasite interaction which shared a resource. Their key result was that the invasion condition for a mutant in such interacting populations could be captured in simple, biologically interpretable expressions (analagous to our equation 2).

Day et al. (2007) developed a similar model for studying two interacting populations, applying it to the specific case of a mutualism between corals and zooxanthellae. Their model also allowed explicit tracking of gene frequencies in both populations.

We adopted Van Baalen and Jansen (2001)'s general approach and notation, but developed a biological model for simple replicators. We studied the case of one population acting as a cooperator and the other acting to affect association and dissociation rates. We argue that this is relevant to molecular replicators, but it is also of potential interest to models of interacting populations in general: because when the association rates can evolve, this will affect evolutionary dynamics.

We used an approach developed by Hurford et al. (2009) to derive the invasion conditions. This is why our invasion condition takes a different form than Van Baalen and Jansen (2001); we found that the Next Generation Theorem allowed us to pull out terms defining the invasion condition that were more easily interpretable from a biological standpoint in this context.

We also extended Van Baalen and Jansen (2001)'s approach in the appendix by: (i) tracking same-type pairs, and therefore allowing for interactions within populations (Section S2), and (ii) explicitly tracking the offspring of complexes, allowing these subpopulations to have distinct dynamics from the larger population (Section S3).

Table S1: Summary of key notation

Notation	Definition
$[X]$	Density of replicator type X
$[Y]$	Density of replicator type Y
$[XY]$	Density of replicator pairs
ρ_i	Total production of type i replicators on their own
θ_i	Total production of type i replicators from pairs
r_i	Baseline replication rate of type i replicator
μ_i	Rate of destruction of type i replicators
β	Baseline association rate
δ	Baseline dissociation rate
k	Degree of density dependence
T	Total density of replicators in system
η	Density dependent replication, $= 1 - k[T]$
r_{ij}	Baseline rate at which ij pairs immediately pair again
κ	Byproduct benefit to being in pair
c	Degree of association trait
d	Degree of cooperative enzymatic activity trait
ζ	Increase in association rate due to association mutation
ξ	Decrease in dissociation rate due to association mutation
ω	Increase in replication rate of type Y due to cooperative enzymatic mutation
λ	Increase in the rate pairs re-pair due to cooperative enzymatic mutation
α	Increase in the rate pairs re-pair due to association mutation

References

Day, T., Nagel, L., Oppen, M. J. v., and Caley, M. J. (2007). Factors affecting the evolution of bleaching resistance in corals. *The American Naturalist*, 171(2):E72–E88.

Dieckmann, U. and Law, R. (1996). The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of mathematical biology*, 34(5-6):579–612.

Geritz, S. A., Metz, J. A., Kisdi, É., and Meszéna, G. (1997). Dynamics of adaptation and evolutionary branching. *Physical Review Letters*, 78(10):2024.

Hurford, A., Cownden, D., and Day, T. (2009). Next-generation tools for evolutionary invasion analyses. *Journal of the Royal Society Interface*, 7(45):561–571.

Law, R. and Dieckmann, U. (1998). Symbiosis through exploitation and the merger of lineages in evolution. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1402):1245–1253.

Metz, J. A., Nisbet, R. M., and Geritz, S. A. (1992). How should we define 'fitness' for general ecological scenarios? *Trends in Ecology & Evolution*, 7(6):198–202.

Rand, D. A., Wilson, H., and McGlade, J. M. (1994). Dynamics and evolution: evolutionarily stable attractors, invasion exponents and phenotype dynamics. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 343(1305):261–283.

Van Baalen, M. and Jansen, V. A. (2001). Dangerous liaisons: the ecology of private interest and common good. *Oikos*, 95(2):211–224.

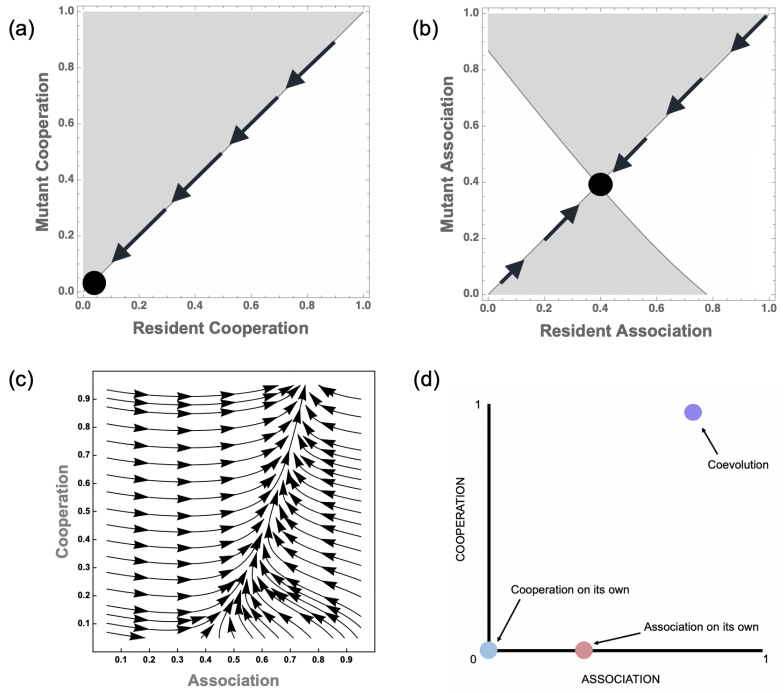

Figure S1: The coevolution of enzymatic activity and association when allowing for same-type pairs to form. Grey shaded areas show regions of state space where selection is negative, and white areas where selection is positive. Arrows depict the direction of evolution in state space along neutral lines. Evolutionarily stable strategies are depicted by solid circles. (a) The evolution of enzymatic activity in X is not favoured. In the absence of association, cooperative enzymatic activity cannot evolve. (b) The evolution of association in Y. In the absence of cooperative enzymatic activity in X, some intermediate level of association in Y is favoured. (c) The coevolution of cooperative enzymatic activity and association. Arrows depict the direction of selection in both traits at a given point in state space. When traits are allowed to coevolve, cooperative enzymatic activity and association both evolve from anywhere in state space, with association reaching higher values than in (b), and enzymatic activity evolving towards its maximal value of 1. (d) A schematic of (a)-(c). Solid circles depict the evolutionarily stable resting point of both populations depending on whether each population evolves independently or evolve jointly. Coevolution favours higher values in both traits. Values for parameters in (a)-(d) are: $\alpha = 20, \lambda = 20, \zeta = 5, \xi = 1, \omega = 20, r_Y = 2.3, r_X = 2.1, r_{XY} = 0.9, r_{XX} = 0.01, r_{YY} = 0.01, k = 0.01, \mu_Y = 1.1, \mu_X = 1.1, \kappa = 100, \beta = 0.01, \delta = 0.9$. All figures generated graphically from the equations described in the Supplementary Material using Mathematica Software version 11.3.0.0.

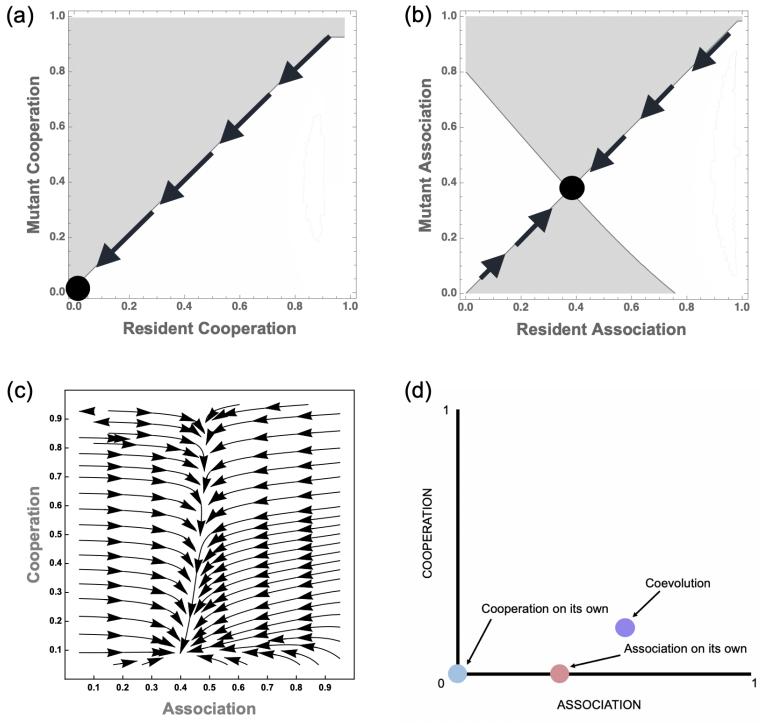

Figure S2: The coevolution of enzymatic activity and association in an explicit model of pairing. Grey shaded areas show regions of state space where selection is negative, and white areas where selection is positive. Arrows depict the direction of evolution in state space along neutral lines. Evolutionarily stable strategies are depicted by solid circles. (a) The evolution of enzymatic activity in X is not favoured. In the absence of association, cooperative enzymatic activity cannot evolve. (b) The evolution of association in Y. In the absence of cooperative enzymatic activity in X, some intermediate level of association in Y is favoured. (c) The coevolution of cooperative enzymatic activity and association. Arrows depict the direction of selection in both traits at a given point in state space. When traits are allowed to coevolve, cooperative enzymatic activity and association both evolve from anywhere in state space, with both traits reaching higher values than in (b) and (c). (d) A schematic of (a)-(c). Solid circles depict the evolutionarily stable resting point of both populations depending on whether each population evolves independently or evolve jointly. Coevolution favours higher values in both traits. Values for parameters in (a)-(d) are: $\alpha = 20, \lambda = 20, \zeta = 5, \xi = 1, \omega = 20, r_Y = 2.3, r_X = 2.1, r_{XY} = 0.9, k = 0.01, \mu_Y = 1.1, \mu_X = 1.1, \kappa = 100, \beta = 0.01, \delta = 0.9, \psi = 1$. All figures generated graphically from the equations described in the Supplementary Material using Mathematica Software version 11.3.0.0.